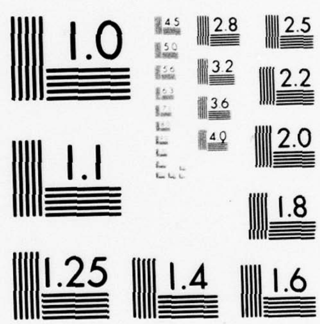
AD-A072 849 AIR FORCE GEOPHYSICS LAB HANSCOM AFB MA F/G 4/2
DESIGN STUDY FOR THE DEVELOPMENT AND USE OF MODEL OUTPUT STATIS--ETC(U)
DEC 78 H R GLAHN, K F HEBENSTREIT
UNCLASSIFIED AFGL-TR-79-0002 NL

| OF |

AD
A072849

END
DATE
FILMED
9-79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFGL-TR-79-0002

DESIGN STUDY FOR THE DEVELOPMENT AND USE OF MODEL OUTPUT STATISTICS IN AUTOMATED AVIATION WEATHER FORECASTING

Harry R. Glahn
Karl F. Hebenstreit

Techniques Development Laboratory
Systems Development Office
8060 - 13th St.
Silver Spring, MD 20910

Final Report
1 April 1976 - 30 September 1978

1 December 1978

Approved for public release; distribution unlimited

AIR FORCE GEOPHYSICS LABORATORY
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
HANSCOM AFB, MASSACHUSETTS 01731

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

14

1. REPORT NUMBER: AFGL-TR-79-0002

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

9

6

4. TITLE (and Subtitle): DESIGN STUDY FOR THE DEVELOPMENT AND USE OF MODEL OUTPUT STATISTICS IN AUTOMATED AVIATION WEATHER FORECASTING.

5. TYPE OF REPORT & PERIOD COVERED: Final Report. 1 Apr [redacted]76 - 30 Sep [redacted]78

6. PERFORMING ORG. REPORT NUMBER

10

7. AUTHOR(s): Harry R. Glahn [redacted] Karl F. Hebenstreit

8. CONTRACT OR GRANT NUMBER(s): ESD 8-0839

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
Techniques Development Laboratory
Systems Development Office
8060 - 13th St., Silver Spring, Md. 20910

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62101F 667010AA

17 10

16

11. CONTROLLING OFFICE NAME AND ADDRESS:
Air Force Geophysics Laboratory
Hanscom AFG, MA 01731
Contract Monitor: Donald A. Chisholm/LYU

11

12. REPORT DATE: 1 Dec [redacted]78

13. NUMBER OF PAGES: 37 pages

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

12 37p.

15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Statistical-dynamical weather prediction; Model output statistics; Weather impact indicator; Linear regression; Regression estimation of event probabilities; Transnormalized regression probability; Product moment correlation; Tetrachoric correlation; Biserial correlation; Brier score

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Air Weather Service is preparing for the WWMCCS era by developing a support concept whereby weather information inputs will be in terms of Weather Impact Indicators (WIIs) which specify the probability that all stages of a mission will have "favorable" weather. This report presents the design of a system based on the proven Model Output Statistics (MOS) approach to provide, in real-time, acceptably skillful objective estimates of WII probabilities. The system is mission oriented allowing the user to specify the limiting values for the weather

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

409578

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Item 20: Continued

elements which impact his mission. If forecast equations for such a mission profile are not already available in the system, forecast equations will be developed, applied, and archived for future use. Acceptable accuracy of these forecasts is essential. A user will not be satisfied with probabilities which, through poor reliability, lead to unexpected losses or which, through poor resolution, furnish little discrimination over and above a simple application of climatology. The design analysis for a real-time MOS system for global application shows such a system to be feasible within the resources of a large-scale computing center such as that operated by the National Oceanic and Atmospheric Administration or the Air Force Global Weather Central. A series of experiments was performed which tested a ceiling-category forecast system based on a specialized version of regression called transnormalized regression probability (TRP). The results of the TRP system were compared with results of the operational system based on the regression estimation of event probabilities (REEP) technique. In our tests, the TRP model did not perform as well as the operational equations based on REEP. However, the TRP model performance was significantly better than climatology and approached the performance of the REEP model at the 24-h projection.

Accession For	
NTIS GRA&I	✓
DDC TAB	
Unannounced	
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or special
A	

PREFACE

The design and coding of software to build the predictand and predictor data bases used in this study was provided by Frank Lewis of GE/MATSCO. John Neander, also of GE/MATSCO, assisted in the running of the forecast and verification programs. The authors gratefully acknowledge their contribution to the study.

CONTENTS

FIGURES

TABLES

DESIGN STUDY FOR THE DEVELOPMENT AND USE OF MODEL OUTPUT STATISTICS IN AUTOMATED AVIATION WEATHER FORECASTING

1. INTRODUCTION

The advent of the automated Worldwide Military Command and Control System (WWMCCS) will alter the way in which weather information is provided to operational decision makers within the Department of Defense. Optimal strategies for committing forces will be sought by examining computer simulations of a wide range of options subject to varying operational and enviromental constraints. These simulations will require inputs which assess in probabilistic terms the existence of specified future conditions of the atmosphere. The decision maker will be guided by simulation output which will order the various options in terms of the overall probability of mission success. The Air Weather Service is preparing for the WWMCCS era by developing a support concept whereby weather information inputs will be in terms of Weather Impact Indicators (WII's) which specify the probability that all stages of a mission will have "favorable" weather. The threshold values which define "favorable" weather will vary from mission to mission and will be provided by the operator as constraints affecting the proposed mission.

The Techniques Development Laboratory (TDL) of the National Weather Service has continued to refine an innovation in statistical-dynamical weather prediction called Model Output Statistics (MOS). This technique provides probability forecasts for a large number of weather variables by developing the statistical relationships between parameters forecast by operational numerical prediction models and the resultant occurrence of weather (Glahn, 1976). The MOS methodology has reached a high state of development with a solid record of success in each new application.

Developing a system to provide real-time and acceptably skillful objective estimates of WII probabilities for a large number of missions is a tremendous challenge--one that has not been previously addressed in enough depth for one to be able to say the best way (or perhaps even a workable way) of going about meeting it! The various considerations in designing a forecast WII (FWII) system fall into two categories: accuracy of forecasts and practicality of use.

Acceptable accuracy of forecasts is a must. A user will not long be satisfied with probabilities which, through poor reliability, lead to unexpected losses or which, through poor resolution, furnish little discrimination over and above a simple application of climatology.

Also, the system must be able to produce WII's when, and in the quantity, needed by users. Many requests will require rapid (on the order of minutes)

responses. Although there will also be many "routine" WII's needed, a capability must be developed and maintained in a high state of readiness to produce upon demand.

In this report, we (1) describe and discuss a possible design for an automated FWII system and show results of test runs which produce FWII's, and (2) describe the results of a program which assessed the accuracy of various numerical-statistical techniques, including a specialized version of regression called transnormalized regression probability (TRP).

2. THE DESIGN OF AN FWII SYSTEM

We have within TDL rather sophisticated developmental (Glahn, 1974) and operational systems for producing objective forecasts with regression or logit models based on the output from numerical models, observations, and climatology. We evaluate about 50,000 equations and produce over 250,000 forecasts daily. Fig. 1 shows the flow of data through the MOS developmental programs. However, our system does not provide for rapid response to new requirements, nor is it designed for use outside the United States. We use the system in making forecasts at points where past observations are available for a developmental sample, although some generalization is possible to other locations. A real-time FWII system can be designed along the lines shown in Fig. 2. Each portion of the figure is identified by a letter and discussed below using that letter as an index. References to computer capabilities are in terms of the IBM 360/195, and especially the system operated by the National Oceanic and Atmospheric Administration (NOAA). Based on this analysis we conclude that such a real-time system is feasible and practical using "off-the-shelf" technology.

a. Analysis Program

This would actually be a group of programs employing different statistical models such as regression, logit, or TRP. Data input, CP (central processor) time, and core requirements would vary considerably depending on the model and other considerations. However, such a program could usually be engineered to run in 500K bytes (1 byte = 8 bits) of core and use about 1 minute of CP time.

b. Historical Model Output

This is the primary predictor set. Model forecasts would exist at a large number of gridpoints and interpolation to predictand locations would be made by a module of the analysis program. If a grid of 1977 points (the NMC octogon) were saved for each of 194 fields (the TDL 0000 GMT collection, which includes forecasts out to 168 hrs.) and a datum occupied 16 bits (the present NMC procedure), one model run would require 194,194 32-bit words (including 12 identifiers per field). Therefore, one IBM 3330-1 disk would hold data from about 127 runs. It would require 5 disks to hold a 635-case sample, which is about the size sample presently used at TDL. Another five disks would be required for the 1200 GMT cycle. This is not an untenable

MOS DEVELOPMENT SYSTEM

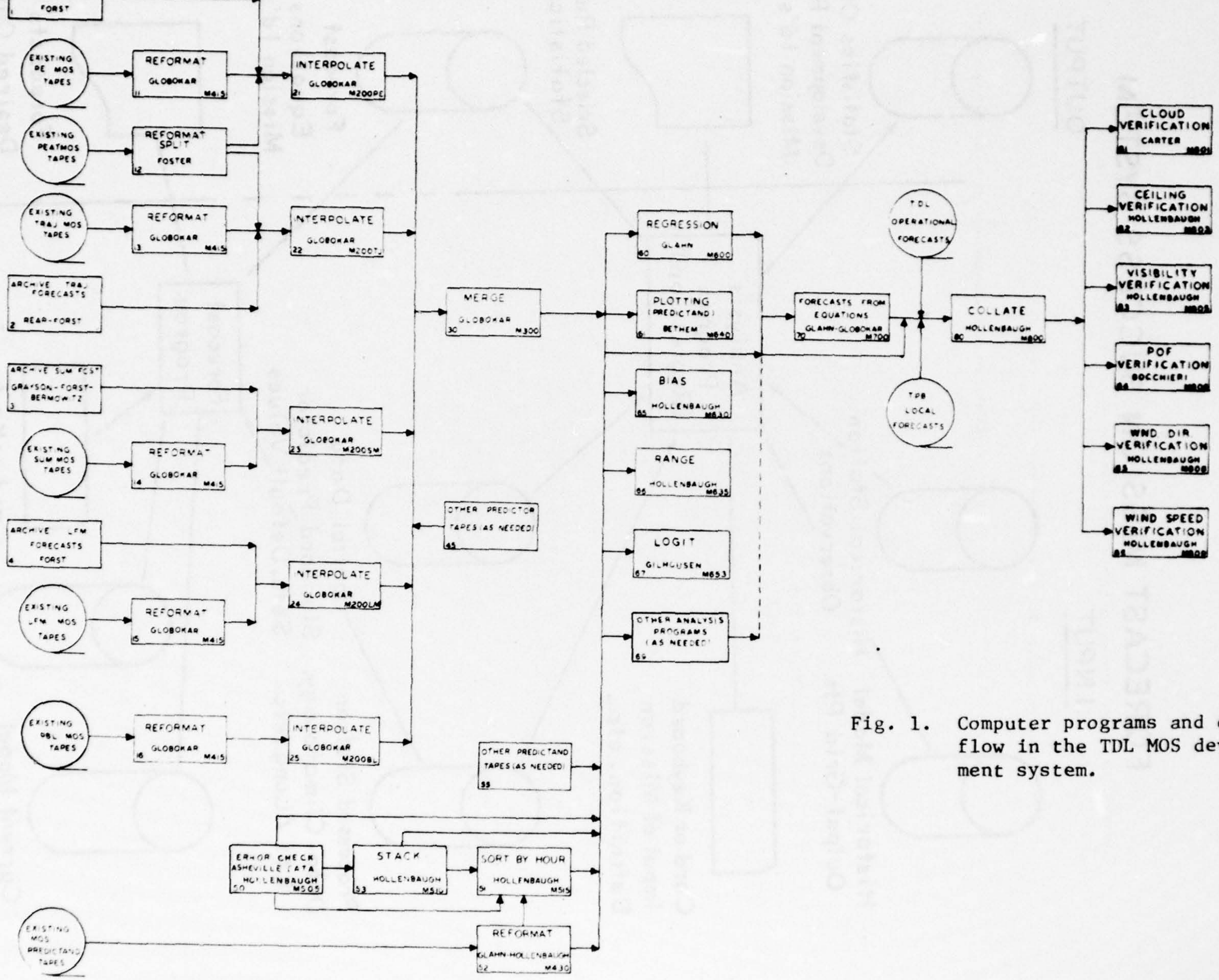


Fig. 1. Computer programs and data flow in the TDL MOS development system.

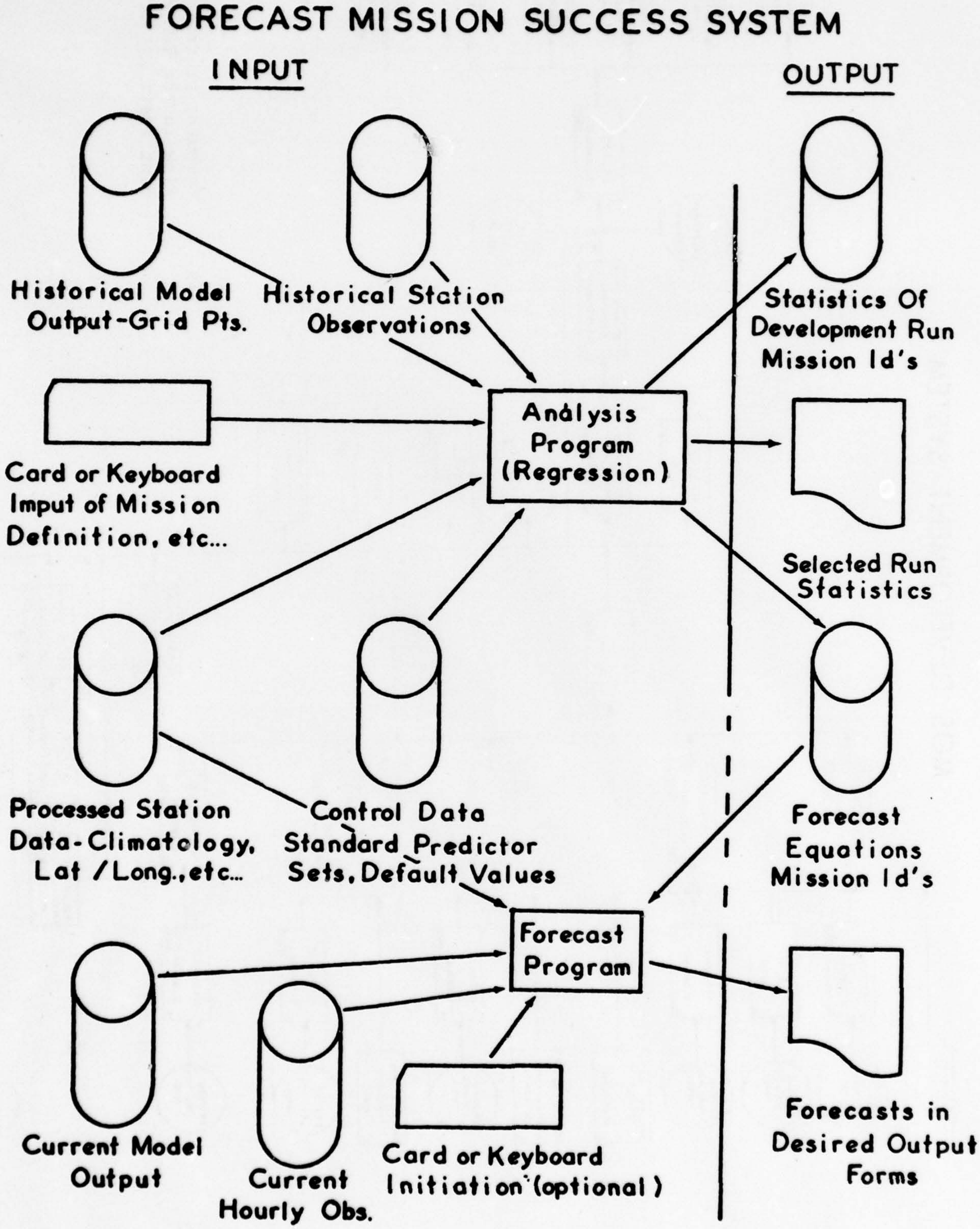


Fig. 2. Possible structure of a real-time FWII system.

number of disks--the NOAA 360/195 systems currently have 64 on-line spindles and upgraded IBM 3330-11 diskpacks which hold twice as much data as the IBM 3330-1's.

c. Historical Station Observations

This is the predictand set. Observations would exist for specific locations. TDL's collection of 255 stations, 8 observations per day of each of 18 variables, occupies 2,027,640 words per 6-month (183 days) season. Therefore, one disk would hold 12.2 seasons or a sample of 2232 days. If an FWII collection were to contain 1000 stations, 24 observations per day for each of 20 variables, one disk would hold about 170 days of data. Therefore, 3.8 3330-1 disks (or 1.9 3330-11's) would hold a 635-day sample.

Our format is designed specifically for our MOS system and is highly efficient for that purpose. The data are optimally packed and are arranged synoptically. If a collection is to cover a large portion of the globe, it may be more advantageous to have subsamples covering areas about the size of the North American Continent with some overlap of neighboring areas. With this data organization, analysis programs would be more efficient in the sense that only portions of the complete data set would have to be read on any given run. This same concept could, and probably should, be applied to the model output data set discussed above.

d. Processed Station Data

This file would contain values for the same locations as does the predictand data set. Typically, this would be station elevation, station location in terms of latitude and longitude and also in terms of the numerical grid, and climatology of the various elements. Probably no more than 0.1 disk would be required for these data.

e. Control Data

Although certain control parameters would be required at run time through card or keyboard, other control data would remain constant for particular predictands. For instance, the predictor variables to screen for ceiling prediction would be contained here and used for each run involving ceiling unless this feature were overridden through cards or keyboard. Only a small portion of a disk would be required for this file.

f. Card or Keyboard Initiation of Analysis Program

The run would be initiated by operator request and necessary information provided through card input or console keyboard.

g. Statistics of Development Run

Various statistics and control information of the development run would be archived according to a set of mission I.D.'s, such as reduction of variance (if appropriate to the model being used), sample used, predictand

definition, beginning and ending dates (month and day) for which equations are valid, etc. Information contained here would completely identify the run and could be used to initiate a rerun.

h. Selected Run Statistics

Enough information would be printed to allow the operator to judge that the run completed successfully and that the results can be used by the forecast program.

i. Forecast Equations

These are the equations that would be used by the forecast program and permanently archived for possible future use. The same mission I.D.'s would be used as in the statistics file. Associated with each set of equations would be a date. This date would initially be the date the equations were generated; each time the equations were used, the existing date would be replaced by the current date. About once a month a program would purge all equations and associated statistics that had not been used for, say, 12 months. One disk would hold equations and associated statistics for about 68,000 missions and associated submissions. This is probably sufficient for a practical-sized file.

j. Forecast Program

This is the program that would make the operational forecasts. It would be initiated by the console operator to produce a forecast for which forecast equations already exist; it could be initiated upon completion of the analysis program; or it could be initiated on a routine basis to provide forecasts for routine requirements. Quite likely, initiation could best be accomplished by automatically providing and updating through the console or analysis program a list of forecasts needed during the next few days. The computer would periodically scan this list to see if any forecasts were needed that had not been made for which the input data were available.

k. Current Model Output

Current model output would be provided in the same format as the historical model output. This output would be migrated to the historical file within 24 hours of its generation or if convenient, could actually be the same physical file as the historical data.

l. Current Hourly Observations

Current observations would be provided for input to forecast equations. It is possible this file could be the same as that containing the historical station observations; however, it is likely that the historical file would be added to once a day after careful quality control of the data, rather than added to each time data are received.

m. Card or Keyboard Initiation of Forecast Program

The forecast program could be initiated by the console operator. This might best be done by updating the list of forecasts needed. Then, the forecast program would discover these forecasts were needed on its next scan.

n. Forecasts

Finally, the needed forecasts would be made available to the user in the form desired. This would include printout for quality control by the duty meteorologist.

3. EXAMPLE FWII'S

The main analysis program used in the TDL MOS system is the M600 Regression Program (Glahn, et al., 1975). This program is quite general in many respects, but does assume that all predictors apply at the predictand point. That is, even though data from several stations can be combined to develop generalized equations, each predictand station has predictors that apply specifically to it. Since a mission may involve several stations and different predictand variables at those stations, M600 could not be used without modification. Also, more versatility in defining the predictand was necessary.

Therefore, an extensive modification of M600 produced M601 which was then used to run two examples. As an added feature, M601 will also make FWII's on the dependent data sample. The two examples are discussed below.

In both cases the majority of possible predictors were from 1200 GMT National Meteorological Center's runs of the Limited-area Fine Mesh (LFM) model (Howcroft and Desmarais, 1971; Gerrity, 1977). Observations at 1500 GMT (tau=3) were also used. The 570-case sample consisted of the 4 winter seasons 1972-73 through 1975-76.

a. FWII Example No. 1

The mission was defined to be:

- Ceiling $\geq$ 400 ft at Washington, D.C. (DCA) at a forecast projection (tau) = 15 hours from cycle time.
- Ceiling $\geq$ 400 ft at Omaha, Neb. (OMA) at tau = 18 hours
- Ceiling $\geq$ 400 ft at Washington, D.C. (DCA) at tau = 24 hours.

Predictors were valid at both Washington and Omaha. In addition to the probability of success of the entire mission, the probability of each "element" of the mission is also produced. Each element and the total mission has its own equation, but all use the same predictors. No assumption is made about independence of predictors or predictands. Unfortunately, because of the finite sample-overfitting problem, some sets of probabilities

are not completely consistent; this will be pointed out in the discussion of specific estimates below.

Table 1 shows FWII's for a few days in the dependent sample (last column). Also shown are the probabilities of each element of the mission. Climatology for the dependent sample is shown on the last line. In the lower left of each box containing a probability is either an "S" representing a success of that element or mission or an "F" representing a failure, i.e., the specified condition was or was not satisfied on the particular day and hour. Finally, in the last column, the number in parenthesis the probability obtained by multiplying together all element probabilities. We can note the following:

(1) Some estimates of success exceed 100%--element (2) on 10/28/72 and on 11/22/72. This is a characteristic of the regression model--estimates are not restricted to the 0 to 100% interval. Another model, such as logit, would eliminate this problem. (In practice, this is seldom a problem.)

(2) The failures generally occur with low probabilities of success. However, since these are only a few selected examples from the dependent data set, no conclusions should be drawn about accuracy.

(3) Climatological probability of success is rather high for each element.

(4) The most noteworthy conclusion that can be drawn from Table 1 is that the product of the individual element probabilities is very close to the mission probability, indicating the elements are relatively independent. In the cases shown, the products are slightly less than the mission probabilities as would be expected by slight dependence.

b. FWII Example No. 2

The total mission was defined to be:

- Ceiling $\geq$ 100 ft and visibility $\geq$ 1/2 mi at Washington, D,C, at tau = 15 hours

- Ceiling $\geq$ 500 ft and visibility $\geq$ 1 mi at Washington, D.C. at tau = 24 hours

- Target 1 clouds $\leq$ .2 coverage at tau = 18 hours

- Target 2 clouds $\leq$ .2 coverage at tau = 18 hours

Target 1 clouds $\leq$ .2 coverage was defined to be $\leq$ .2 coverage at all three stations Aberdeen, Fargo, and Huron, S.D. Target 2 was comprised of San Antonio, Austin, and Victoria, Tex. From Table 2 we can note:

Table 1. FWII's for a few days in the dependent sample for Example No. 1.

Date (12 GMT)	(1) 15-Hr. DCA Cig$\geq$ 400Ft.	(2) 18-Hr. Omaha Cig$\geq$400 Ft.	(3) 24-Hr. DCA Cig$\geq$400Ft.	(1)(2)(3)
10/28/72	S 80	S 102	F 78	F 69 (64)
11/01/72	S 87	S 85	S 69	S 53 (51)
11/19/72	F 67	S 99	S 100	F 67 (66)
11/22/72	S 93	S 103	S 73	S 72 (70)
02/01/73	F 76	S 91	S 76	F 54 (53)
12/12/75	S 100	F 74	S 90	F 69 (67)
Climatology	96	96	94	88 (87)

Table 2. FWII's for a few days in the dependent sample for Example No. 2.

Date (12 GMT)	(1) 15-Hr. DCA Cig ≥ 100 Ft Vis ≥ 1/2 Mi.	(2) 24-Hr. DCA Cig ≥ 500 Ft Vis ≥ 1 Mi.	(3) Target 1 ≤ .2 Clouds	(4) Target 2 ≤ .2 Clouds	(1)(2)(3)	(1)(2)(4)	(1)(2)(3)(4)
11/13/72	S 98	F 63	F 55	S 64	F 49	F 43	F 28 (22)
11/15/72	S 101	S 99	F 10	F 26	F 13	F 28	F 05 (02)
12/02/72	S 101	S 97	F 05	S 26	F 06	S 28	F 02 (01)
12/26/72	S 100	S 87	S 57	S 104	S 59	S 94	S 54 (52)
02/01/73	F 94	S 86	F 44	S 67	F 41	F 56	F 21 (24)
03/14/73	S 101	S 96	F 06	F 09	F 07	F 11	F 05 (01)
11/28/73	S 99	S 103	S 38	S 94	S 40	S 93	S 40 (36)
11/13/74	S 100	S 94	F 05	S 53	F 04	S 50	F -01 (02)
Climatology	99	94	27	36	26	34	09 (09)

(1) The climatologies of mission elements (1) and (2) are rather high, but those of elements (3) and (4) are much lower.

(2) The success of the complete mission has a rather low probability on most days; some of the cases shown such as 12/26/72 is much higher than usual. The climatological probability of success is only 9%.

(3) The chance of success of a partial mission (columns 6 and 7) is usually much higher than of a complete mission.

(4) The four elements of the mission are quite independent as evidenced by the fact that the product of the four element probabilities is only very slightly lower than the mission success probability determined directly by equation.

(5) Some inconsistencies are present. For instance, the probability of success of element (2) is greater than 100% on 11/28/73. Also, the probability of success of elements (1), (2), and (3) is higher than the probabilities of success of the single element (3) on 12/2/72.

4. COMPARISON OF NUMERICAL-STATISTICAL FORECAST TECHNIQUES

This section presents the results of a series of experiments which tested a forecast system based on a specialized version of regression called transnormalization regression probability (TRP) (Boehm, 1976). TRP has several unique features. First, preprocessing of predictand and predictor data is required. In this step the data are transformed and normalized (hence, the term coined by Boehm, op. cit., transnormalized) using cumulative distribution functions (cdfs) for each predictor and the predictand conditioned on the time of day (TOD). Second, biserial and tetrachoric correlation in the place of product-moment correlation are recommended for variables characterized by "spiked" cdfs. For example, at many locations the "no ceiling/unlimited-ceiling" cloud condition has a frequency of occurrence of more than 0.4. Under this condition the last datum point for ceiling on the cdf curve has a cumulative frequency of about 0.6 with the rest of the cases (approximately 40% of the sample) assigned to the "no ceiling/unlimited-ceiling" category. Third, TRP uses a prediction equation based on an extension of the bivariate normal equation (Gringorten, 1972) to the multivariate case, a procedure which assumes that the resultant distribution will be miltivariate normal.

4.1 Description of Experiments to Test the TRP Model

A hierachy of experiments was performed to assess the effect of varying parameters associated with the processing of the predictand and predictor data. A 6-category cloud ceiling predictand, stratified as shown in Table 3, was used for these experiments. The variable parameters included the period of record (POR) of the data, the station grouping, and the specificity of the climatology. Varying the parameters changed the data base used

to produce the predictor and predictand cdf's from which transnormalized values were obtained for the various experiments.

Table 3. Category limits for ceiling.

Category
$0 \leq 1 < 200$ ft
200 ft $\leq 2 <$ 500 ft
500 ft $\leq 3 <$ 1000 ft
1000 ft $\leq 4 <$ 3000 ft
3000 ft $\leq 5 \leq$ 7500 ft
7500 ft ≤ 6 (or no ceiling)

For the most part, the existing and unique TDL data base was used. A 4-year LFM model cool season (October-March) data set was used to build model predictor cumulative distribution functions. Four projection lengths were considered in the study, namely 6-, 12-, 18-, and 24-h from the 0000 GMT cycle. Approximately 40 predictors were available for each projection. These were either fields produced by the LFM model (e.g. boundary layer relative humidity) or predictors derived from the model fields such as 850-mb temperature-dewpoint spread. For example, Table 4 contains the list available at the 24-h projection. The values of each of these candidate predictors were interpolated to 233 NWS station locations producing an interpolated predictor tape of 233 predictor values for about 160 predictors for approximately 600 days. These data were further processed to obtain the required cdf for each predictor and were ultimately transnormalized before being offered to the screening regression program.

Predictand data consisting of observed values of ceiling height for the 233 stations were available for seven winters beginning in October 1969. For 9 stations in the mid-Atlantic region these predictand data were augmented with data beginning in 1949 yielding an approximately 25-year POR. This allowed us to run an experiment to test the sensitivity of the forecast results to the period of record used as the basis for transnormalizing the predictand data. Predictand cdfs were obtained from a 4-year, a 7-year, and, in the case of the 9 stations, a 25-year data base. Transnormalized 0300 GMT predictand data were offered as predictors to the screening regression program. The use of an observed ceiling condition as a predictor conforms to current NWS practice.

The processing software written to transform the "raw" predictor and predictand data was general enough to build the required cdfs for either a single station or for a group of stations (region), and for a month or for an entire 6-month season. For both the predictand and predictors, cdfs were conditioned on the time of day with separate curves for 0600, 1200, 1800, 2400 GMT. Since the 0300 GMT predictand values were used as a predictor, a cdf for the 0300 GMT ceiling condition was also required.

Table 4. List of LFM model predictors available at the 24-h projection.

Predictor Type		Projection
LM BL RH		24
LM BL RH	S5	24
LM BL RH DEL12S5		24
LM BL RH DEL24S5		24
LM L1 RH		24
LM L1 RH	S5	24
LM MEAN RH		24
LM MEAN RH	S5	24
LM MRH DELT12S5		24
LM P WATER		24
LM P WATER	S5	24
LM P AMT	S5	24
LM BL U	S5	24
LM BL V	S5	24
LM 850 U	S5	24
LM 850 V	S5	24
LM 700 U		24
LM 700 V		24
LM 500 U		24
LM 500 V		24
LM 700 WND SPS5		24
LM 500 WND SPS5		24
LM BL VV	S5	24
LM 850 VV	S5	24
LM 700 VV	S5	24
LM 1000 HGT	S5	24
LM 850 HGT	S5	24
LM 500 HGT		24
LM 500 HGT	S5	24
LM 8.5-10 TH	S5	24
LM 7-10 TH	S5	24
LM 5-10 TH	S5	24
LM BL POT T		24
LM BL POT T	S5	24
LM 10T-10DP	S5	24
LM 8.5T-8.5DP	S5	24
LM TT INDEX	S5	24
LM K INDEX	S5	24
LM G INDEX	S5	24

LM = LFM Model
BL = Boundary Layer
RH = Relative humidity
S5 = 5-point smoothing
VV = Vertical Velocity
TH = Thickness
DEL = Time change
HGT = Height

The cdfs for predictand data were constructed by accumulating a count of the number of cases for each reportable value of ceiling height. A total of 93 classes were possible in the scheme we used. The values from 1 to 50 correspond to ceilings in the range of 0 to 4900 feet at intervals of 100 feet; 51 to 60 correspond to ceilings in the range 5000 to 9500 feet at intervals of 500 feet; 61 to 85 correspond to ceilings in the range of 10,000 to 34,000 feet at intervals of 1000 ft, 86 to 92 were used for ceilings between 35,000 and 65,000 feet at intervals of 5000 feet. All those observations above 65,000 feet and the "no ceiling" cases were assigned the value 93. Of course, the count in many of the "boxes" in the scheme was zero when producing a given cdf, but this scheme did insure that all possible ceiling observations were assigned unambiguiously to one of the 93 possible classes.

The cdfs for each LFM model predictor were constructed by dividing the range of values of the predictor (at a location) in the dependent sample into 100 equally spaced intervals. The cdf for each predictor, for each station, for each time of day was then derived by assigning values of the predictor in the dependent sample to its proper interval, and finally, accumulating the data. In the case of cdfs for regions the predictor values were assigned to 200 equally spaced intervals, counted, and accumulated as with the station cdfs. The increase in the number of intervals allowed us to handle a wider range of values which exist over a region without a corresponding increase in the absolute value of the width of an interval.

While 3 transnormalized methods are proposed by Boehm, op. cit., only the look-up table method was used in this test. The transnormalizing methods not considered include parameteric and polynomial curve fitting and the modeling of predictand and predictor distributions. The look-up table method was chosen for testing because it is a straight-forward technique which remains as close as possible to the data. The other methods, which only approximate the cdfs derived from the data, introduce error requiring judgements as to the adequacy of the curves fitted to the data. Handling these judgements would further complicate the interpretation of the test results.

Each interval used to construct a cdf curve corresponds to a value or range of values of the predictand or predictors. The value of the cdf corresponding to the interval can be mapped onto the normal ogive to obtain the equivalent normal deviate (END) corresponding to the "raw" value of the variable. The rank-ordered cdf data were used to construct look-up tables which provided the END of each predictor and the predictand given the TOD and the integer assigned to the interval associated with the value of the variable. These lookup tables were than used to convert the predictand and predictor data in the dependent data set into the corresponding END values for input to the screening regression program. Linear interpolation was used to assign values of END to predictor datum values which didn't correspond exactly to the datum value assigned to a given interval in the lookup table. Independent data predictor values which fall outside the range of values in the dependent sample were arbitrarily assigned ENDS of the first or last intervals of the lookup tables. Such data correspond to values of the predictor more than 4 standard deviations from the mean of the dependent sample.

The TRP technique was examined by conducting 3 experiments which considered single station vs. regional grouping and seasonal vs. monthly data sets. In the first experiment, called TRP1, the predictand and predictor lookup tables used to convert raw data to END values were based on seasonal sets for stations grouped into 21 regions previously defined in related technique development studies. Next we transformed the raw data using seasonal data for each station (TRP2). Finally, lookup tables based on data for single station during a month (e.g. four Octobers) were used (TRP3). This final step resulted in a profusion of lookup tables, i.e. one table for each 233 stations, for each of 6 months, for each of 40 predictors, for each 4 projections--almost a quarter of a million tables when those required for the predictand are also included in the total. Table 5 contains a summary of the three experiments along with a listing of the factors varied in each experiment.

Once the transnormalized data tapes for the dependent predictand and predictor sets were prepared, the TDL screening regression program was used to derive forecast equations. The predictand and predictors were offered as continuous variables and 12-term regression equations developed which estimate the END of the forecast, $\hat{m}=a_1x_1+\ldots+a_{12}x_{12}$. The forecast $\hat{m}$ can be thought of as a correction (based on forecast LFM fields and the 0300 GMT surface observations) to be applied to the climatic value of the predictand $\hat{y}$ using the predictive equation of the TRP model:

$$Pr(Y < y|x_1,X_2\ldots X_{12}) = \Phi\left[\frac{\hat{y} - \hat{m}}{\sqrt{1 - R^2}}\right]$$

i.e. the probability that the predictand Y will be less than the breakpoint y given the values $x_1,\ldots x_{12}$ of the predictors is given by normal cumulative probability of being less than $\frac{\hat{y} - \hat{m}}{\sqrt{1 - R^2}}$ where:

$\hat{y}$ = the END of the category breakpoint y obtained from the climatic lookup table for the TOD for which the predictor is desired.

$\hat{m}$ = the estimate of the forecast from the regression equation valid as the TOD for which the prediction is desired.

R = the multiple correlation coefficient.

Forecast equation sets were obtained in each of the 3 experiments for 4 forecast projections. In each case, the forecast equations (regression coefficients and multiple correlation coefficient) were obtained for groups of stations (21 regions) using the transnormalized data for the entire dependent data set. This procedure produced an equation to be applied at each station in a region on each day in the dependent and independent data samples. The pooling of the data in this manner provided stable forecast

Table 5. Hierarchy of TRP model experiments for ceiling prediction. The portion of the experiment that is changed from the previous experiment is underlined.

	Experiment	Characteristics	Comment
TRP1	Predictand	4 yr/region-season lookup	closest comparison to operational TDL MOS
	Predictor	4 yr/region-season lookup	
	Regression Coefficients	region/season	
TRP2	Predictand	4 yr/single station-season lookup	
	Predictor	4 yr/single station-season lookup	
	Regression Coefficients	region/season	
TRP3	Predictand	7 yr/single station-monthly-lookup	
	Predictor	4 yr/single station-monthly-lookup	
	Regression Coefficients	region/season	

equations. It is especially justifiable in the case of TRP since the forecast procedure explicitly adjusts the event climatology for a given station (TRP2 and TRP3) based on forecast model output.

The TRP model yields equations that forecast the cumulative probability of each ceiling category defined in Table 3. The cumulative probabilities were converted to exclusive category probabilities for comparison with the operational MOS (REEP) system. The equations developed from the dependent sample were applied to dependent and independent data for 233 stations for 06-, 12-, 18-, and 24-h projections for each experiment. The dependent sample contained 594 cases while the independent sample contained 178 cases.

The probability forecasts generated by TRP and REEP for each of the ceiling categories were verified using the Brier score (Brier, 1950). Because the Brier score is a function of the sample climatology it varies from location to location and from year to year. As a control, a climatic Brier score was obtained by verifying a forecast of the seasonal relative frequency of each of the categories for each station for each day.

A separate experiment was conducted to test an aspect of TRP proposed by Boehm (op. cit.). He suggests that the accuracy of the TRP model forecasts should increase as the representativeness of the predictand climatic cdfs is improved. In other words, the longest possible period of record should be used. To assess this assertion, monthly predictand lookup tables based on an average of 25 years of ceiling observations for each of 9 stations listed in Table 6 were constructed and the data transnormalized. Our objective was to compare forecast results to those obtained using the 7 year predictand sample for the same stations. Equations were developed using each predictand set (called TRP7 and TRP25) and forecasts for this group of stations were made for both the dependent and independent samples. The Brier score was used as the measure of accuracy of the various models with the Brier score for climatology used as a normalizing standard for comparison.

4.2 Results of the Test of the TRP Model

It is not meaningful to compare the equations produced for REEP and TRP by the screening regression program. A REEP equation directly forecasts the probability of occurrence of an exclusive ceiling category. In the TRP model the regression equation provides an estimate of the END of the ceiling height based on model predictors. This value is substituted in the TRP model equation which, after transforming from END to a probability, provides the cumulative probability that the ceiling will be less than a given breakpoint. A subtraction must then be performed to obtain the probability of the exclusive category. Therefore, in this test we compare the final results of the forecast process for each of the candidate models.

Tables 7 through 14 summarize the Brier score for each ceiling category for each of the forecast techniques when applied to the dependent and independent data sets. The results are presented for 4 projections, 6-, 12-, 18-, 24-, from the 0000 GMT cycle time. The column headed CLIM is the Brier score achieved by using the seasonal relative frequency of the category for each station as the forecast probability vector. The column headed REEP

Table 6. Stations used to test the impact of using long-term predictand climatic data.

Raleigh-Durham, NC
Philadelphia, PA
Richmond, VA
Washington-National, DC
Wilmington, DE
Allentown, PA
Harrisburg, PA
Baltimore, MD
Washington-Dulles, VA

Table 7. Brier score for each ceiling category for climatology (CLIM), TDL operational REEP equations (REEP), transnormalized regression probability (TRP) equations, experiment 1 (TRP1), experiment 2 (TRP2), and experiment 3 (TRP3). Dependent data, 4 winters, 1972-1976. Projection 6 hours from 0000 GMT cycle.

Category	CLIM	REEP	TRP1	TRP2	TRP3
1	.02928	.02124	.02449	.02440	.02475
2	.06452	.04941	.05277	.05212	.05193
3	.09598	.07521	.08143	.08059	.08037
4	.21549	.15698	.17528	.17227	.17180
5	.22194	.17433	.19557	.19237	.19182
6	.42702	.18743	.20524	.20330	.20583
All*	.52712	.33230	.36739	.36253	.36325

* The Brier score for all categories (i.e. for the ceiling) is one-half the sum of the scores for the individual categories.

Table 8. Same as Table 7 except for projection 12 hours from 0000 GMT cycle.

Category	CLIM	REEP	TRP1	TRP2	TRP3
1	.04629	.04240	.04805	.04766	.03127
2	.08604	.07633	.07875	.07775	.05779
3	.11219	.10008	.10302	.10218	.08723
4	.23178	.20215	.21242	.20884	.19329
5	.21703	.19667	.20940	.20629	.20926
6	.44114	.25867	.27934	.27571	.26408
All	.56724	.43815	.46549	.45922	.42146

Table 9. Same as Table 7 except for projection 18 hours from 0000 GMT cycle.

Category	CLIM	REEP	TRP1	TRP2	TRP3
1	.01184	.01130	.01219	.01247	.01210
2	.05931	.05506	.05608	.05586	.05439
3	.11037	.09882	.10006	.09933	.09800
4	.26106	.22961	.23707	.23479	.23139
5	.20821	.19577	.20291	.20024	.20926
6	.44135	.27982	.29380	.29764	.28910
All	.54607	.43519	.45106	.45017	.44712

Table 10. Same as Table 7 except for projection 24 hours from 0000 GMT cycle.

Category	CLIM	REEP	TRP1	TRP2	TRP3
1	.01117	.01093	.01171	.01157	.01150
2	.04931	.04575	.04640	.04522	.04498
3	.08432	.07688	.07795	.07704	.07622
4	.21177	.18750	.19264	.18974	.18742
5	.23042	.21481	.22146	.21804	.21678
6	.41602	.27982	.28775	.28454	.28286
All	.50151	.40785	.41896	.41308	.40988

Table 11. Same as Table 7 except independent data, 1 winter, 1976-1977.

Category	CLIM	REEP	TRP1	TRP2	TRP3
1	.02241	.01680	.01883	.01924	.02030
2	.04688	.03730	.03979	.03990	.04057
3	.07263	.05766	.06178	.06186	.06264
4	.18219	.13376	.14677	.14559	.14643
5	.20746	.16044	.17967	.17868	.17869
6	.39974	.17176	.18918	.18823	.19193
All	.46566	.28886	.31801	.31675	.32028

Table 12. Same as Table 8 except independent data, 1 winter, 1976-1977.

Category	CLIM	REEP	TRP1	TRP2	TRP3
1	.03560	.03268	.03591	.03588	.03710
2	.06494	.05791	.05924	.05818	.05935
3	.08542	.07671	.07839	.07830	.07847
4	.20005	.17361	.18181	.17989	.17798
5	.20911	.18672	.19925	.19782	.19760
6	.42276	.24618	.26636	.26453	.26682
All	.50894	.38691	.41048	.40730	.40866

Table 13. Same as Table 9 except independent data, 1 winter, 1976-1977.

Category	CLIM	REEP	TRP1	TRP2	TRP3
1	.01106	.01111	.01186	.01172	.01201
2	.04532	.04315	.04405	.04330	.04377
3	.08422	.07485	.07543	.07539	.07613
4	.21830	.18918	.19324	.19360	.19166
5	.19935	.18500	.19194	.19122	.18922
6	.41547	.26298	.27491	.28284	.27474
All	.48686	.38314	.39572	.39904	.39377

Table 14. Same as Table 10 except independent data, 1 winter, 1976-1977.

Category	CLIM	REEP	TRP1	TRP2	TRP3
1	.03336	.00805	.00887	.00874	.00853
2	.06038	.03443	.03445	.03461	.03494
3	.07771	.06069	.06139	.06143	.06193
4	.18179	.15560	.15820	.15707	.15798
5	.19478	.20349	.20914	.20797	.20683
6	.28670	.26463	.26949	.26764	.27077
All	.41736	.36345	.37077	.36813	.37049

gives the Brier score obtained from the probability vector provided by the TDL operational equations. The columns headed TRP1, TRP2, and TRP3 give the Brier scores for the three experiments which tested the TRP model. Again, the differences are in the data input to the lookup tables used to trans-normalize the raw predictand and predictor values: TRP1 used the total dependent sample and combined 233 stations into 21 regions to produce regional-seasonal tables; TRP2 used the total dependent sample for each station to produce single station-seasonal tables; TRP3 used monthly data (e.g. data from 4 Octobers) for each station to produce single station-monthly tables.

For the United States as a whole, REEP generally outperformed TRP in both the dependent and independent samples. Out of 72 pairwise comparisons of the individual categories using the Brier score, REEP was better in all but 9 in the dependent sample and all 72 in the independent sample.

As would be expected, when the data were stratified into smaller sub-samples in building the lookup tables (progressing from TRP1 to TRP3), the Brier scores, in general, decreased (became better) on dependent data. However, in many cases, this decrease did not hold up on independent data; of the 24 pairwise comparisons of TRP1 and TRP3 in Tables 11 through 14, TRP3 was better in 11, and for the 24 pairwise comparisons of TRP2 and TRP3, TRP3 was better in only 7. (These are comparisons of individual category scores, not including the category <u>All</u>.)

Table 15 summarizes the percent improvement over climatology for all categories combined for the various experiments. With exception of the 12-h projection on the dependent data, REEP has the largest percentage improvement on both the dependent and independent data. The advantage of REEP is largest at the shorter projections. By 24-h, there is little difference between TRP (in its various forms) and REEP. While the TRP model has larger Brier scores than REEP, clearly its improvement over climatology is substantial.

With regard to the 25.7 percent improvement for TRP3 on the dependent sample at the 12-h projection, we point out that it is possible to have the dependent sample Brier score show a dramatic improvement over a long term climatology when the dependent sample is very small. This comes about because of the serial correlation in the observations. For example, if a larger than usual number of low conditions occurs in a small dependent sample, the equivalent normal deviate will reflect a higher than usual probability. Comparing the sharpness of such probabilities against those derived from larger samples will favor the probabilities based on the smaller sample. However, an independent sample comparison will not be so influenced and will, in fact, favor the scheme based on the larger sample when all other things are held constant.

The results of the separate experiment where TRP based on a 25-year predictand climatology (TRP25) was compared to TRP based on a 7-year predictand climatology (TRP7) are given in Tables 16 through 23 for both the dependent and independent samples. TRP7 was better on the dependent sample for all 24 pairwise comparisons except one. This would be expected, since the 4-year predictor dependent sample represents a larger percentage of the

Table 15. Percent improvement over climatology.

Projection (h)	Dependent (4-winters 72-76) REEP	TRP1	TRP2	TRP3	Independent (1-winter 76-77) REEP	TRP1	TRP2	TRP3
6	37.0	30.3	31.2	31.1	38.0	31.7	32.0	31.2
12	22.8	17.9	19.0	25.7	24.0	19.3	20.0	19.7
18	20.3	17.4	17.6	18.1	21.3	18.7	18.0	19.1
24	18.7	16.5	17.6	18.3	12.9	11.2	11.7	11.2

Table 16. Brier score for each ceiling category for forecasts for nine stations in the mid-Atlantic region for TDL operational equations (REEP), for transnormalized regression probability equations using monthly predictand station lookup tables based on 7 years of data (TRP7), and tables based on 25 years of predictand data (TRP25). Dependent data, 4 winters, 1972-1976. Projection 6 hours from 0000 GMT cycle.

Category	REEP	TRP7	TRP25
1	.02227	.02631	.02683
2	.05993	.06821	.06895
3	.06839	.07377	.07426
4	.12845	.13989	.14192
5	.21308	.23042	.23070
6	.20657	.21383	.23144
All*	.34935	.37622	.38705

* The Brier score for all categories (i.e. for the ceiling) is one-half the sum of the scores for the individual categories.

Table 17. Same as Table 16 except for 12-h projection.

Category	REEP	TRP7	TRP25
1	.03339	.04042	.04181
2	.09214	.09425	.09669
3	.09225	.09653	.09705
4	.14919	.14900	.15120
5	.24979	.26010	.26211
6	.26090	.27642	.27985
All	.43883	.45836	.46436

Table 18. Same as Table 16 except for 18-h projection.

Category	REEP	TRP7	TRP25
1	.00727	.00787	.00777
2	.05472	.05398	.05455
3	.10067	.10104	.10269
4	.17020	.16634	.16841
5	.25467	.25886	.26190
6	.27043	.27737	.28316
All	.42898	.43273	.43924

Table 19. Same as Table 16 except for 24-h projection.

Category	REEP	TRP7	TRP25
1	.00914	.01011	.01080
2	.05760	.05846	.05872
3	.07913	.08052	.08111
4	.14621	.14377	.14635
5	.23601	.23887	.24158
6	.26258	.26461	.26781
All	.39534	.39817	.40319

Table 20. Same as Table 16 except for independent data, one winter 1976-1977.

Category	REEP	TRP7	TRP25
1	.02172	.01425	.01448
2	.04439	.04501	.04308
3	.05342	.05697	.05660
4	.09067	.08679	.08590
5	.20443	.20369	.20175
6	.19634	.19417	.19258
All	.30549	.30044	.29720

Table 21. Same as Table 17 except for independent data, one winter 1976-1977.

Category	REEP	TRP7	TRP25
1	.01810	.02090	.02234
2	.06552	.06321	.06242
3	.07627	.08078	.08017
4	.10322	.10029	.10079
5	.22630	.22376	.22331
6	.23004	.23383	.23355
All	.35973	.36139	.36129

Table 22. Same as Table 18 except for independent data, one winter, 1976-1977.

Category	REEP	TRP7	TRP25
1	.00142	.00192	.00192
2	.03982	.04019	.03905
3	.07210	.06870	.06770
4	.11065	.10788	.10556
5	.24717	.24488	.24226
6	.25913	.26418	.26699
All	.36515	.36388	.36174

Table 23. Same as Table 19 except for independent data, one winter 1976-1977.

Category	REEP	TRP7	TRP25
1	.00501	.00601	.00637
2	.03730	.03918	.03805
3	.05534	.06026	.05928
4	.09035	.09304	.09190
5	.22813	.22783	.22734
6	.25551	.24998	.25136
All	.33582	.33815	.33715

7-year predictand data base than of the 25-year data base (the 4-year predictor sample was contained in both the 7- and 25-year samples). However, when applied to the independent data TRP25 was better than TRP7 in 17 out of 24 comparisons. This indicates that long term climatology provides a somewhat better basis for transnormalizing the data, at least in this case where the subsample is rather small to start with. For these 9 stations which are a subset of the 233 station sample, TRP25 was slightly better than REEP yielding better Brier scores in 14 of the 24 later comparisons. The REEP equations do not account for station climatic difference, while the TRP model used in this experiment explicitly uses transnormalized data based on monthly-station lookup tables. Techniques which account for individual station bias in REEP regionalized equations would improve the performance of the REEP model.

4.3 Tetrachoric Correlation Test

For predictors and predictands with "spiked" cumulative distributions, Boehm (op. cit.) proposes the subsititution of the tetrachoric correlation for the product moment correlation in the correlation matrix for those cases where both variables are dichotomized and biserial correlation for cases where one variable is dichotomized and the other continuous. This procedure can be criticized on principle. The so-called normal equations, which provide the regression solution, are based on product moment considerations. To substitute the tetrachoric or biserial correlation may introduce uncertain and dubious outcomes.

Consider three variables with pairwise correlations r_{12}, r_{13}, and r_{23}. These correlations, all computed from the same data base, must satisfy the consistency relationship:

$$r_{12}^2 + r_{23}^2 + r_{13}^2 - 2r_{12}r_{23}r_{13} \leq 1.$$

One cannot arbitrarily substitute other estimates for r_{12}, r_{13}, and r_{23} without the danger of violating this relationship. Tetrachoric and biserial correlation are approximations to the product moment correlation that can be calculated when the basic data are not available in continuous form. Further, when these approximations were introduced into the statistical literature, their use in the correlation matrix to solve regression equations was not proposed.

To demonstrate the problems which can arise using the suggested procedure, a series of experiments was performed with 5 observed ceiling predictors used to forecast the ceiling at 1200 GMT, i.e., observed ceilings at 0900, 0600, 0300, 0000, and 2100 GMT to forecast the ceiling at 1200 GMT. The END's for both the predictors and the predictand were obtained from single station-monthly lookup tables. Breakpoints to dichotomize the predictors and predictand were selected at the median END, the median minus one-half the standard deviation, and the median minus the standard deviation. Using these highly correlated predictors a problem developed which had not occurred with the product moment correlations. The correlation matrix would not invert for 2 of the 21 regions when the breakpoint of the median END minus one standard deviation was used. When matrix inversion cannot be achieved, the computerized regression procedure is aborted and prediction equations can not be obtained. The predictors used in the MOS technique may be highly correlated. If tetrachoric correlation is substituted, the consistency relation may or may not be satisfied; if it is not, a "blow-up" of the linear regression solution will result.

5. Conclusions

With the large-scale computer technology available today, a real-time MOS system to provide Weather Impact Indicators is feasible. Because of the volume of data required, much care must be given to organizing predictand and predictor data sets to insure efficient operational application.

In our tests of the TRP model it did not perform as well as the TDL operational equations based on REEP. Equations based on data which were transnormalized by using monthly-single station lookup tables performed better than seasonal-single station equations on the dependent data set. However, this performance did not hold up on independent data. Equations based on single station predictand/predictor lookup tables were superior to those based on regional tables. For the data which we used in ceiling forecasting, the longer period of record (approximately 25 years) predictand data produced Brier scores on independent data which were a slight improvement over those using the 7-year data set.

The substitution of tetrachoric correlations is questioned since their use will cause problems in obtaining regression equations when the consistency constraints on correlations entering the correlation matrix are violated.

The TRP model performance was signficantly better than climatology and approached the performance of the REEP model at the 24-h projection.

The most serious drawback to the TRP system is the volume of calculations needed to prepare the lookup tables. In addition, the storage capacity required may preclude the use of lookup tables in a real-time operational system. This could be overcome by the use of fitted parametric curves. However, this may introduce additional error and loss of performance, and is not recommended.

TRP does have the advantage of giving a probability of <u>any</u> range of the predictand (e.g. ceiling between 500 and 2000 ft) without prior knowledge of the range needed. Quite likely, however, the use of several categories in REEP and the fitting of a curve to the forecast cumulative probabilities will give as good or better results than TRP and be much less effort.

REFERENCES

Boehm, A. R., 1976: Transnormalized regression probability. AWS-TR-75-259, Air Weather Service (MAS), 52 pp.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. Mon. Wea. Rev., 78, 1-3.

Gerrity, J. F., 1977: The LFM model - 1976: A documentation. NOAA Tech. Memo. NWS NMC-60, National Weather Service, National Oceanic and Atmospheric Administration, U.S. Dept. of Commerce, 60 pp.

Glahn, H. R., 1974: The TDL MOS development system IBM 360/195 version. TDL Office Note 74-14, National Weather Service, National Oceanic and Atmospheric Administration, U.S. Dept. of Commerce, approx. 66 pp.

_____, G. W. Hollenbaugh, and F. T. Globokar (editors), 1975: Computer programs for the MOS development system IBM 360/195 version. TDL Office Note 75-2, National Weather Service, National Oceanic and Atmospheric Administration, U.S. Dept. of Commerce, approx. 200 pp.

_____, 1976: Automated forecasts of surface weather elements. Proceedings of the Seventh Technical Exchange Conference, El Paso, Tex.

Gringorten, I. I., 1972: Conditional probability for an exact, noncategorical initial condition. Mon. Wea. Rev., 100, 796-798.

Howcroft, J. and A. Desmarais, 1971: The Limited-area Fine Mesh (LFM) model. NWS Tech. Proc. Bull., No. 67, National Weather Service, National Oceanic and Atmospheric Administration, U.S. Dept. of Commerce, 11 pp.